



## RAW SEQUENCE LISTING

DATE: 04/16/2002

PATENT APPLICATION: US/09/461,537

TIME: 16:03:11

Input Set : N:\Crf3\RULE60\09461537.raw

Output Set: N:\CRF3\04162002\I461537.raw

## SEQUENCE LISTING

## 3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Royer, John C

6 Moyer, Donna L

7 Yoder, Wendy T

8 Shuster, Jeffrey R

10 (ii) TITLE OF INVENTION: NON-TOXIC, NON-TOXIGENIC, NON-PATHOGENIC

11 FUSARIUM EXPRESSION SYSTEM

13 (iii) NUMBER OF SEQUENCES: 16

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: Novo Nordisk of North America, Inc.

17 (B) STREET: 405 Lexington Avenue, 64th Floor

18 (C) CITY: New York

19 (D) STATE: New York

20 (E) COUNTRY: USA

21 (F) ZIP: 10174-6401

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Floppy disk

25 (B) COMPUTER: IBM PC compatible

26 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

27 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

29 (vi) CURRENT APPLICATION DATA:

C--&gt; 30 (A) APPLICATION NUMBER: US/09/461,537

C--&gt; 31 (B) FILING DATE: 15-Dec-1999

32 (C) CLASSIFICATION:

35 (vii) PRIOR APPLICATION DATA:

36 (A) APPLICATION NUMBER: 08/816,915

37 (B) FILING DATE: 13-MAR-1997

41 (viii) ATTORNEY/AGENT INFORMATION:

42 (A) NAME: Agris Dr., Cheryl H.

43 (B) REGISTRATION NUMBER: 34,086

44 (C) REFERENCE/DOCKET NUMBER: 4216.240-US

46 (ix) TELECOMMUNICATION INFORMATION:

47 (A) TELEPHONE: 212-867-0123

48 (B) TELEFAX: 212-878-9655

51 (2) INFORMATION FOR SEQ ID NO: 1:

53 (i) SEQUENCE CHARACTERISTICS:

54 (A) LENGTH: 30 base pairs

55 (B) TYPE: nucleic acid

56 (C) STRANDEDNESS: single

57 (D) TOPOLOGY: linear

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

61 TCGGATCCA TGGTCAAGTT CGCTTCCGTC

ENTERED

30

## RAW SEQUENCE LISTING

DATE: 04/16/2002

PATENT APPLICATION: US/09/461,537

TIME: 16:03:11

Input Set : N:\Crf3\RULE60\09461537.raw

Output Set: N:\CRF3\04162002\I461537.raw

```

64 (2) INFORMATION FOR SEQ ID NO: 2:
66   (i) SEQUENCE CHARACTERISTICS:
67       (A) LENGTH: 30 base pairs
68       (B) TYPE: nucleic acid
69       (C) STRANDEDNESS: single
70       (D) TOPOLOGY: linear
72   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
74 GACCTCGAGT TAAGCATAGG TGTCAATGAA                               30
77 (2) INFORMATION FOR SEQ ID NO: 3:
79   (i) SEQUENCE CHARACTERISTICS:
80       (A) LENGTH: 998 base pairs
81       (B) TYPE: nucleic acid
82       (C) STRANDEDNESS: single
83       (D) TOPOLOGY: linear
85   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
87 ATCATCAACC ACTCTTCACT CTTCAACTCT CCTCTCTTGG ATATCTATCT CTTCAACCATG       60
89 GTCAAGTTCG CTTCGTCGT TGCACCTGTT GCTCCCCTGG CTGCTGCCGC TCCTCAGGAG       120
91 ATCCCAACA TTGTTGGTGG CACTTCTGCC AGCGCTGGCG ACTTTCCTT CATCGTGAGC       180
93 ATTAGCCGCA ACGGTGGCCC CTGGTGTGGA GGTTCCTCTCC TCAACGCCAA CACCGTCTTG       240
95 ACTGCTGCCC ACTGCGTTTC CGGATACGCT CAGAGCGGTT TCCAGATTCG TGCTGGCAGT       300
97 CTGTCTCGCA CTTCTGGTGG TATTACCTCC TCGCTTTCCT CCGTCAGAGT TCACCCTAGC       360
99 TACAGCGGAA ACAACAACGA TCTTGCTATT CTGAAGCTCT CTACTTCCAT CCCCTCCGGC       420
101 GGAAACATCG GCTATGCTCG CCTGGCTGCT TCCGGCTCTG ACCCTGTGCG TGGATCTTCT       480
103 GCCACTGTTG CTGGCTGGGG CGCTACCTCT GAGGGCGGCA GCTCTACTCC CGTCAACCTT       540
105 CTGAAGGTTA CTGTCCCTAT CGTCTCTCGT GCTACCTGCC GAGCTCAGTA CGGCACCTCC       600
107 GCCATACCA ACCAGATGTT CTGTGCTGGT GTTCTCTCCG GTGGCAAGGA CTCTTGCCAG       660
109 GGTGACAGCG GCGGCCCCAT CGTCGACAGC TCCAACACTC TTATCGGTGC TGTCTCTTGG       720
111 GGTAACGAT GTGCCCAGCC CAACTACTCT GGTGTCTATG CCAGCGTTGG TGCTCTCCGC       780
113 TCTTTCATTG ACACCTATGC TTAAATACCT TGTTGGAAGC GTCGAGATGT TCCTTGAATA       840
115 TTCTCTAGCT TGAGTCTTGG ATACGAAACC TGTTTGAGAA ATAGGTTTCA ACGAGTTAAG       900
117 AAGATATGAG TTGATTTTCA TTGGATCTTA GTCCTGGTTG CTCGTAATAG AGCAATCTAG       960
119 ATAGCCCAAA TTGAATATGA AATTGATGA AAATATTC                       998
122 (2) INFORMATION FOR SEQ ID NO: 4:
124   (i) SEQUENCE CHARACTERISTICS:
125       (A) LENGTH: 248 amino acids
126       (B) TYPE: amino acid
127       (C) STRANDEDNESS: single
128       (D) TOPOLOGY: linear
130   (ix) FEATURE:
131       (A) NAME/KEY: Protein
132       (B) LOCATION: 1..224
134   (ix) FEATURE:
135       (A) NAME/KEY: Peptide
136       (B) LOCATION: -24..0
137       (D) OTHER INFORMATION: /product= "OTHER"
138 /note= "Label=pre-propeptide"
140   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
142   Met Val Lys Phe Ala Ser Val Val Ala Leu Val Ala Pro Leu Ala Ala
143                               -20                               -15                               -10

```

## RAW SEQUENCE LISTING

DATE: 04/16/2002

PATENT APPLICATION: US/09/461,537

TIME: 16:03:11

Input Set : N:\Crf3\RULE60\09461537.raw

Output Set: N:\CRF3\04162002\I461537.raw

```

145   Ala Ala Pro Gln Glu Ile Pro Asn Ile Val Gly Gly Thr Ser Ala Ser
146           -5                      1                      5
148   Ala Gly Asp Phe Pro Phe Ile Val Ser Ile Ser Arg Asn Gly Gly Pro
149       10                      15                      20
151   Trp Cys Gly Gly Ser Leu Leu Asn Ala Asn Thr Val Leu Thr Ala Ala
152       25                      30                      35                      40
154   His Cys Val Ser Gly Tyr Ala Gln Ser Gly Phe Gln Ile Arg Ala Gly
155           45                      50                      55
157   Ser Leu Ser Arg Thr Ser Gly Gly Ile Thr Ser Ser Leu Ser Ser Val
158           60                      65                      70
160   Arg Val His Pro Ser Tyr Ser Gly Asn Asn Asn Asp Leu Ala Ile Leu
161           75                      80                      85
163   Lys Leu Ser Thr Ser Ile Pro Ser Gly Gly Asn Ile Gly Tyr Ala Arg
164       90                      95                      100
166   Leu Ala Ala Ser Gly Ser Asp Pro Val Ala Gly Ser Ser Ala Thr Val
167       105                      110                      115                      120
169   Ala Gly Trp Gly Ala Thr Ser Glu Gly Gly Ser Ser Thr Pro Val Asn
170           125                      130                      135
172   Leu Leu Lys Val Thr Val Pro Ile Val Ser Arg Ala Thr Cys Arg Ala
173           140                      145                      150
175   Gln Tyr Gly Thr Ser Ala Ile Thr Asn Gln Met Phe Cys Ala Gly Val
176           155                      160                      165
178   Ser Ser Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Ile
179       170                      175                      180
181   Val Asp Ser Ser Asn Thr Leu Ile Gly Ala Val Ser Trp Gly Asn Gly
182       185                      190                      195                      200
184   Cys Ala Arg Pro Asn Tyr Ser Gly Val Tyr Ala Ser Val Gly Ala Leu
185           205                      210                      215
187   Arg Ser Phe Ile Asp Thr Tyr Ala
188           220

```

191 (2) INFORMATION FOR SEQ ID NO: 5:

193 (i) SEQUENCE CHARACTERISTICS:

194 (A) LENGTH: 1206 base pairs

195 (B) TYPE: nucleic acid

196 (C) STRANDEDNESS: single

197 (D) TOPOLOGY: linear

199 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

201 GAATTCTTAC AAACCTTCAA CAGTGGAGAC TTCCGACACG ACATATCGAT CCTTTGAAGA      60
203 TACGGTGAGC GTCAGATCAT GAATTCATA CATCCTCACG TCCTTCCTCT TTCAAACAT      120
205 GCAAAGTCCT TCTAGTACCT CCCAAACTT GATTACGCG CTCTCCAATC AAAAGTACCT      180
207 TCCAAAAGTG ATCTACCTCA GCTCTAGATC AGGGCACCTA TTCGCAAAGA TCTACAAGCT      240
209 GAACTAGTAA GCATAGCGGG AGAATATCCC ACATCATTCG AGAAGGCCTT CGTATTAGAC      300
211 CTAGTGGGAT CGACAGAAAA GATAAGACGG AGATAGATGC TATGTTTGGA AGGTAGGGGA      360
213 TGGAATAGGA TGCAACAGGT ATTGGCATAA GCGATGCAAT AGGTGCATCT AGAACTAGG      420
215 TGACAGACTG GCCACAGAGG TGTATCCTAT GCAGGTCGAT GCGTGCGTTA TCGCAGGGCT      480
217 GCTATTGCGT GGTGGTGGCT ACAAAGTTT TATGTGGTTT CCAGTTTCAG AATATTGGGC      540
219 CATTGTGATT GATGGCGCAT GACCGAATTA TAGCAGTGAA CCCCGCCAG AGTAGTAGTG      600
221 CAGATGCGCT TTGATGCTTG GCGATTCTC GGGCTAAATA ACTCCGTTG GTCTGTAGAA      660
223 TGCTGACGCG ATGATCCTTC GGCATTAATC GTAGATCTTG GGGGGGGATA AGCCGATCAA      720

```

## RAW SEQUENCE LISTING

DATE: 04/16/2002

PATENT APPLICATION: US/09/461,537

TIME: 16:03:11

Input Set : N:\Crf3\RULE60\09461537.raw

Output Set: N:\CRF3\04162002\I461537.raw

```

225 AGACACACTG TAGATCAGCT CTTCGATGAC TCTTACCAGC TTTATAATAA CATTTCATCTT      780
227 GAACGTCTTT TTCGTCCAGT GTTTACCTTT CGTCCTATTT ATCCGTCATA TCCACAGTGT      840
229 TATTGGCGAT AGAGTTATCG ACTTTCCTCA TCGGGATACT GGCCCTGCT GCCAAGGGCC      900
231 TTATATGCCG ATCACTTTCA CGGGAGCATG ATAAGGTAA TGCTTCTTCT GAATGCCGAA      960
233 CTAGACTACG GAACAACGGA GCTTAGTACC AGAAAGGCAG GTACGCCTAT TCGCAAACCTC    1020
235 CGAAGATACA ACCAAGCAAG CTTATCGCGG GATAGTAACC AGAGAGGCAG GTAAGAAGAC    1080
237 ACAACAACAT CCATAGCTAT GTAGATTCTC GAATATAAAA GGACCAAGAT GGACTATTCTG    1140
239 AAGTAGTCTA TCATCAACCA CTCTTCACTC TTCAACTCTC CTCTCTTGGA TATCTATCTC    1200
241 TTCACC                                         1206
244 (2) INFORMATION FOR SEQ ID NO: 6:
246     (i) SEQUENCE CHARACTERISTICS:
247         (A) LENGTH: 1188 base pairs
248         (B) TYPE: nucleic acid
249         (C) STRANDEDNESS: single
250         (D) TOPOLOGY: linear
252     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
254 TAAATACCTT GTTGAAGCG TCGAGATGTT CCTTGAATAT TCTCTAGCTT GAGTCTTGGA      60
256 TACGAAACCT GTTTGAGAAA TAGGTTTCAA CGAGTTAAGA AGATATGAGT TGATTTTCAGT    120
258 TGGATCTTAG TCCTGGTTGC TCGTAATAGA GCAATCTAGA TAGCCCAAAT TGAATATGAA    180
260 ATTTGATGGA AATATTCATT TCGATAGAAG CAACGTGAAA TGTCTAGCAG GACGAAAAGT    240
262 AGATCAAGGC TGTATGTTC CCCGACCAAC CTACCTTGAT GTCAGTCTGC GAGTCGTGTG    300
264 CAGTGACCCA GAATGATGGA TTGACTTGGA CATTCTCTGT CTATGAAGTA TTATGAACAT    360
266 GAATATCGTT TCCTCATTAT CTATGTTGGC AGCCTAAAGT TTTACCATAT AGCTAGCAAT    420
268 CAGTCAAGTA TCTGCGTATG AAGGGTTGTT AAGCCAGGAC GGTATCAGCG TTGAATATTT    480
270 AAAGAATGAT ATGAGATAAT CAACATTGAC ATGATAAAAG AAAAGGGGAA ACAAATTGTG    540
272 CATATAGTAA AGACTTCAGG TCGACCCCTC AATAGACATA TGCGAACCGA AAACCAACAG    600
274 GATACAATTT ATAGATAAGT ATAACACAG TTATCTGTCT GCCGAACAAA TACTCTTTTG    660
276 TGAAACAAAT GAAGAGTACA TAAGCTACAG TTCCTCAGTA GGAACATCCT TTACAATAAC    720
278 TCCCTTGACT TCCTTCAGCT TCTCAATAGC CTCCAAAGTC ATCGGTCTGC CATCAAGGCA    780
280 CGTCAGCTCT GGTGTAGCAT ACAGCAGTGC CATACTTACG GAGGATAGGA AGTGGGAGGA    840
282 ATCGTTCGTG TCTGCCTCCA AAAATCGACA CCAGTGTCCT TTTTGACGAT ACTGATATGG    900
284 TGGTAAGCTT GGGAGTCTAT TGTTGACGTT GCATCACTTA CTTAAGCACG GTTTCATTCC    960
286 TCTGCTGATA GTCCTCCAAC TTCTCGAAGT CGTAAACGAT GGCCTATAGT ATCTTATTGA   1020
288 GAAATATGTC TTCTCAGAAA ATTATATCTT GTTTACCTTT CGGTCCGCCA TGGCTGCTAA   1080
290 AACTGCTGGG AAATTCAAAA GCGCAGCACA AGCAGCAAGA GTGATGGGCA CAACGTGATA   1140
292 TGTTGATAAA AGCATCAGTA TCGATAAGTT CCACTCAGAA ACCTGCAG                    1188
295 (2) INFORMATION FOR SEQ ID NO: 7:
297     (i) SEQUENCE CHARACTERISTICS:
298         (A) LENGTH: 1060 base pairs
299         (B) TYPE: nucleic acid
300         (C) STRANDEDNESS: single
301         (D) TOPOLOGY: linear
303     (ix) FEATURE:
304         (A) NAME/KEY: CDS
305         (B) LOCATION: 10..924
307     (ix) FEATURE:
308         (A) NAME/KEY: mat_peptide
309         (B) LOCATION: 73..924
311     (ix) FEATURE:

```

## RAW SEQUENCE LISTING

DATE: 04/16/2002

PATENT APPLICATION: US/09/461,537

TIME: 16:03:11

Input Set : N:\Crf3\RULE60\09461537.raw

Output Set: N:\CRF3\04162002\I461537.raw

```

312      (A) NAME/KEY: sig_peptide
313      (B) LOCATION: 10..72
315      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
317 GGATCCAAG ATG CGT TCC TCC CCC CTC CTC CCG TCC GCC GTT GTG GCC      48
318      Met Arg Ser Ser Pro Leu Leu Pro Ser Ala Val Val Ala
319      -21 -20      -15      -10
321 GCC CTG CCG GTG TTG GCC CTT GCC GCT GAT GGC AGG TCC ACC CGC TAC      96
322 Ala Leu Pro Val Leu Ala Leu Ala Ala Asp Gly Arg Ser Thr Arg Tyr
323      -5      1      5
325 TGG GAC TGC TGC AAG CCT TCG TGC GGC TGG GCC AAG AAG GCT CCC GTG      144
326 Trp Asp Cys Cys Lys Pro Ser Cys Gly Trp Ala Lys Lys Ala Pro Val
327      10      15      20
329 AAC CAG CCT GTC TTT TCC TGC AAC GCC AAC TTC CAG CGT ATC ACG GAC      192
330 Asn Gln Pro Val Phe Ser Cys Asn Ala Asn Phe Gln Arg Ile Thr Asp
331 25      30      35      40
333 TTC GAC GCC AAG TCC GGC TGC GAG CCG GGC GGT GTC GCC TAC TCG TGC      240
334 Phe Asp Ala Lys Ser Gly Cys Glu Pro Gly Gly Val Ala Tyr Ser Cys
335      45      50      55
337 GCC GAC CAG ACC CCA TGG GCT GTG AAC GAC GAC TTC GCG CTC GGT TTT      288
338 Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asp Phe Ala Leu Gly Phe
339      60      65      70
341 GCT GCC ACC TCT ATT GCC GGC AGC AAT GAG GCG GGC TGG TGC TGC GCC      336
342 Ala Ala Thr Ser Ile Ala Gly Ser Asn Glu Ala Gly Trp Cys Cys Ala
343      75      80      85
345 TGC TAC GAG CTC ACC TTC ACA TCC GGT CCT GTT GCT GGC AAG AAG ATG      384
346 Cys Tyr Glu Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met
347      90      95      100
349 GTC GTC CAG TCC ACC AGC ACT GGC GGT GAT CTT GGC AGC AAC CAC TTC      432
350 Val Val Gln Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe
351 105      110      115      120
353 GAT CTC AAC ATC CCC GGC GGC GGC GTC GGC ATC TTC GAC GGA TGC ACT      480
354 Asp Leu Asn Ile Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Thr
355      125      130      135
357 CCC CAG TTC GGC GGT CTG CCC GGC CAG CGC TAC GGC GGC ATC TCG TCC      528
358 Pro Gln Phe Gly Gly Leu Pro Gly Gln Arg Tyr Gly Gly Ile Ser Ser
359      140      145      150
361 CGC AAC GAG TGC GAT CGG TTC CCC GAC GCC CTC AAG CCC GGC TGC TAC      576
362 Arg Asn Glu Cys Asp Arg Phe Pro Asp Ala Leu Lys Pro Gly Cys Tyr
363      155      160      165
365 TGG CGC TTC GAC TGG TTC AAG AAC GCC GAC AAT CCG AGC TTC AGC TTC      624
366 Trp Arg Phe Asp Trp Phe Lys Asn Ala Asp Asn Pro Ser Phe Ser Phe
367      170      175      180
369 CGT CAG GTC CAG TGC CCA GCC GAG CTC GTC GCT CGC ACC GGA TGC CGC      672
370 Arg Gln Val Gln Cys Pro Ala Glu Leu Val Ala Arg Thr Gly Cys Arg
371 185      190      195      200
373 CGC AAC GAC GAC GGC AAC TTC CCT GCC GTC CAG ATC CCC TCC AGC AGC      720
374 Arg Asn Asp Asp Gly Asn Phe Pro Ala Val Gln Ile Pro Ser Ser Ser
375      205      210      215
377 ACC AGC TCT CCG GTC AAC CAG CCT ACC AGC ACC AGC ACC ACG TCC ACC      768

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/461,537

DATE: 04/16/2002

TIME: 16:03:12

Input Set : N:\Crf3\RULE60\09461537.raw

Output Set: N:\CRF3\04162002\I461537.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]